

THE RELATIONSHIP OF ACADEMIC TASK
DIFFICULTY TO OFF-TASK BEHAVIORS

BY

PAUL J. SIEGEL

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1980

ACKNOWLEDGMENTS

Over the past several years I have had the good fortune of associating with some top-notch professionals in the field of education. I would like to take the opportunity to express my appreciation to a few of these individuals.

Dr. Bill Wolking, my doctoral committee chairman, has been patient, enthusiastic, an expert resource, and dedicated to the learning process. He has provided a model of scholarship which I find worthy of emulation.

I owe a great debt of gratitude to my friend, Dr. Joe Wittmer. Aside from being an excellent educator, powerful administrator, political force to be reckoned with, and humanitarian, Joe has been steadfastly supportive of my professional development.

Thanks go to another doctoral committee member and soul-mate, Dr. Bob Algozzine. Bob's timely advice and thought-sharing has helped a great deal and is appreciated.

Finally, deep appreciation goes to Leila Cantara. Leila typed and proofed my dissertation while readily giving of her superb organizational skills. She exceeded the call of duty.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
ABSTRACT	v
CHAPTER I	
INTRODUCTION	1
Statement of the Purpose	3
Rationale for the Study	4
Limitations	6
Delimitations	6
CHAPTER II	
REVIEW OF THE LITERATURE	7
Relationship of Task Difficulty to Attending Behaviors . . .	7
Interrelationships Among Readability, Achievement, and Behavior	9
Teacher Practices Affecting Classroom Behavior	14
Readability of Academic Materials	17
Inaccuracies of Grade Level Designations in Basal Readers	19
Mismatch of Academic Task and Student Ability	21
Relationship of Ability-Difficulty Matching to Classroom Behavior	24
Summary of Review of Literature	28
CHAPTER III	
METHOD AND PROCEDURES	31
Subjects	31
Setting	31
Materials	31
Dependent Variable	32
Assessment of Student Ability and Level of Material Difficulty	33
Independent Variable	35
Experimental Design	36
Observers	39
Data Collection	40
Data Analysis and Interpretation	41

CHAPTER IV

RESULTS	42
Reliability Measure	43
Effects of Task Difficulty Level on Off-Task Behavior . . .	44
Presentation Order	49
Interest/Satisfaction Rating	51
Correlates of Off-Task Behavior	51
Summary	53

CHAPTER V

DISCUSSION	54
Findings	54
Problems and Limitations of the Study	59
Conclusions and Implications	61
Suggestions for Further Research	64
REFERENCES	66
APPENDIX A DEFINITION OF TERMS	73
B DESCRIPTION OF SUBJECTS	74
C RAW AND SMOOTHED DATA	75
BIOGRAPHICAL SKETCH	80

Abstract of Dissertation Presented to the Graduate
Council of the University of Florida in Partial Fulfillment
of the Requirements for the Degree of Doctor of Philosophy

THE RELATIONSHIP OF ACADEMIC TASK
DIFFICULTY TO OFF-TASK BEHAVIORS

By

Paul J. Siegel

August, 1980

Chairman: William D. Wolking
Major Department: Special Education

This study investigated the relationship between levels of academic task difficulty and time off-task in third grade children. Single subject research methodology was used. A multi-element within sessions design was selected which provided a continuous current baseline to which treatment effects were compared. Data were evaluated by analysis of cumulative duration records and visual inspection of graphical displays.

Five third grade children attending public school in Alachua County, Florida, participated in the study. The reading performance level of each child was assessed by directly measuring reading rate, accuracy, and comprehension. The difficulty level assigned to a particular reader was derived from the child's actual performance with that reader instead of using publisher designated grade levels.

All experimental conditions took place within the regular classroom setting with subjects remaining in their regularly assigned seats. Data were obtained by trained observers who presented subjects with the varying reading tasks during morning class periods regularly scheduled for reading or seatwork. A measure of satisfaction with the material being read was obtained in addition to the primary dependent variable, seconds off-task.

Results of the investigation indicate that (a) difficulty level of academic tasks can be an important factor influencing off-task behavior (for some subjects systematic increases in off-task behavior), (b) subjects tended to dislike the more difficult reading material while enjoying the easier material, and (c) experimenter effects may have a significant influence on the remarkably low levels of off-task behavior demonstrated by some subjects in situations of mismatch.

The results imply that rather than hypothetical or organic causes, factors originating within the classroom environment may often be responsible for off-task behaviors. Frequent direct assessment of student performance level and task difficulty makes it possible to determine whether these undesirable behaviors are classroom based. If variables under teacher control are determined to be the source of the problems, the likelihood of correcting the problems is increased.

CHAPTER I

INTRODUCTION

Foundations of educational thought rest on basic principles and assumptions. Often, these principles and assumptions are easily understood by both theoretician and practitioner alike and thus receive wide general acceptance. This study is concerned with such a frequently assumed general educational principle--the principle that a student's classroom behavior is affected by the difficulty level of his academic materials.

In 1940, Durrell wrote:

The most important factor in meeting a child's reading needs is the provision of reading materials suited to his level of ability. . . . A child with second grade reading ability experiences confusion and discouragement in attempting to read fourth grade books. (p. 65)

This notion of the affects of academic task difficulty has had many other proponents since Durrell (Betts, 1946; Bond & Tinker, 1973; Fry, 1975; Harris & Sipay, 1975; May, 1973). Despite (or because of) such wide general acceptance, the topic has received little research attention. The recent contributions of Jorgenson (1977, 1978) and Jorgenson, Klein, and Kumar (1977) comprise the investigative base for this popular assumption. These studies found that, in general, inappropriate classroom behavior increased as difficulty

level of task material increased. This finding lends a degree of empirical credence to the assumed principle. However, as will be pointed out in the next chapter, research methods used in all three studies leave room for improvement.

It is not difficult for a teacher to overlook the influence of task difficulty on behavior. After all, there is a very wide range of factors, including physical, social, emotional, and intellectual variables which influence classroom behavior. Children coming from broken homes exhibit inappropriate behaviors more frequently than nondeviant children (Bacon, Child, & Barry, 1963), behavior disorders are more common in premature children (Drillen, 1961), and aggressive behavior occurs more often in children whose parents use physical punishment (Becker, 1964). Those factors which cannot be controlled by the classroom teacher, but which are nevertheless considered the cause of the problem, are justifiably thought to be not the teacher's responsibility to correct.

However, educational factors also affect classroom behavior and many of these factors are under direct teacher control. In fact, some classroom teacher practices, such as the inadvertent reinforcement of off-task behaviors (Thomas, Becker, & Armstrong, 1968), inconsistent enforcement of classroom rules (Duke, 1978), and failure to provide needed positive reinforcement for successes or attempts to succeed (Becker, Madsen, Arnold, & Thomas, 1967), actually serve to drive pupils off their assigned tasks and into behaviors considered unacceptable in the classroom. Assigning a child to materials which are much too difficult for him can be another practice, albeit inadvertent, contributing to behavior

problems in the classroom. The classroom teacher possesses a direct approach to remediation of both learning and behavior problems when the relationship between such educational practices and behavior is recognized.

Statement of the Purpose

It is the purpose of this investigation to determine the relationship between varying levels of academic task difficulty and off-task classroom behaviors. Specifically, the question this study will attempt to answer is whether an increase in non-attending behaviors, e.g., off-task glancing, occurs as a result of presenting increasingly difficult reading passages to an elementary school child.

Single-subject research methodology ($N=1$) will be used. This set of procedures emphasizes direct observation of behavior and continuous recording of data over time for one subject. The aim of this methodology is to demonstrate convincingly experimental control between the independent and dependent variables.

During each day of experimental conditions, a subject will read silently for five minutes in three differing readers. The readers will have been shown to have a functionally different difficulty level for each child. A trained observer will record a duration measure of seconds off-task while the child is reading.

The same experiment will be conducted with each of four additional elementary school children. Thus, four complete replications of the study will be available to provide additional evidence of the generalizability of results.

Rationale for the Study

Most classroom teachers are probably aware of the need to provide academic materials of suitable difficulty for their pupils. This is the basic educational principle mentioned earlier. However, these same teachers may be unaware that many of the problem behaviors they cope with daily in their classrooms can be due to a mismatch of ability and task difficulty. This mismatch develops due to a number of factors to be discussed in the next chapter. For now, an example will suffice.

Suppose that a teacher who is cognizant of the wide range of ability within the typical elementary classroom (Goodlad, 1966) needs to decide on which basal reading text to use with each of the children in her¹ class. Relying on the recommendation from the previous year's teacher has some potential usefulness but is also subject to much error (May, 1973). So are standardized reading achievement tests (Harris & Sipay, 1975; Lovitt & Hansen, 1976). If she decides to use a basal reading series with specific grade level designations, she could find herself making very inappropriate matches because these designations are often inaccurate (Britton & Lumpkin, 1977; Mills & Richardson, 1963). The same problem often exists with readability formulae (Bormuth, 1966). And, perhaps the biggest source of error will occur since the teacher usually decides on the instructional level of a pupil at the beginning of a school year and for various reasons seldom changes her

¹For simplicity's sake, this study will use exclusively feminine personal pronouns when referring to classroom teachers.

assessment or placement of him. That is, she seldom reassesses (May, 1973).

It becomes relatively easy for a student to be assigned materials that are much too difficult for his ability. The behaviors manifested by a child in coping with this situation may well be detrimental to the well-being of the individual and the class as a whole.

There are several implications of the relationship between ability/difficulty matching and classroom behavior which point to the need for its study and indicate why the results can be useful. In much special education literature the child with attentional and learning problems is conceptualized as learning disabled. This study hopes to show that it can be the interaction between task and child which is responsible for aberrant behaviors. In this way, the onus of a pathology is removed from the child. Instead of testing the child, emphasis would be placed on testing the appropriateness of the match between the child and his academic assignment.

A second implication of a clear link between specific classroom behaviors and the accuracy of the match is that administrators and teachers may become more sensitive to the need for continual re-evaluation of ability/difficulty matching, especially as new materials and assignments are introduced.

Finally, previous research has shown a positive relationship between attending-to-task behaviors and academic achievement (Cobb, 1972; Lahaderne, 1968; Samuels & Turnure, 1974). Children who attend to task gain most from their instruction while inattention

is significantly related to a lack of school success. It follows that students who are accurately matched to their assignments and therefore attend more, would learn more.

Limitations

The difficulty level of reading materials is determined by many different factors, e.g., concept difficulty, test organization, and reader interest. It would be impossible to hold constant all of these variables when determining a measure of readability.

Delimitations

Only reading has been selected as the academic task for which difficulty levels will be assessed and experimentally manipulated. Whether the findings of this study will generalize to other academic subjects, e.g., math, will have to be demonstrated by other research.

In addition, only third graders in regular public school classrooms have been selected as subjects for this study. Whether similar effects would be obtained with students of more or less academic experience or in special education classes will also have to be demonstrated.

CHAPTER II

REVIEW OF THE LITERATURE

While few research studies have been concerned directly with the relationship between academic task difficulty and classroom behavior (Jorgenson, 1977, 1978; Jorgenson, Klein, & Kumar, 1977), research endeavors in other related fields have contributed significantly to an understanding of the issue. Three general areas of research stand out in particular as major contributors. These include studies dealing with:

1. the relationship of task difficulty to attending behaviors,
2. teacher practices affecting classroom behavior, and
3. the readability of academic materials.

Relationship of Task Difficulty to Attending Behaviors

A popular theoretical assumption since the time of William James (1890) states that appropriate attending behavior is necessary for academic learning to take place. That is, the student who does not attend to an academic task is not likely to learn from it.

However, failing to learn is not the only adverse result of non-attending. Often the behaviors which take the place of attending behaviors, that is off-task behaviors, are inappropriate in a

classroom. Teacher's referrals and school psychologist's reports frequently characterize children with learning and behavior problems as being daydreamers, distractible, and as having a short attention span.

A small body of evidence exists which shows that the complexity of non-academic tasks significantly contributes to the occurrence of off-task behaviors. Keogh, Welles, and Weiss (1972) researched the hypothesis that a difficult and ambiguous puzzle-type task will produce more distractible behaviors than a less difficult task. Four and five-year-old children were presented with simple and complex tasks and their visual orienting behavior or direction of gaze was observed. The children were found to exhibit a considerable amount of glancing away from the complex task but almost no glancing away from the simple task, which was well within their ability repertoire. Other studies (Nottleman & Hill, 1977; Ruble & Nakamura, 1972) have since confirmed that off-task glancing is a sensitive indicant of inattentiveness.

The effects of differential reinforcement and task difficulty on attention to teacher instruction were researched by Bucher and Okovita (1977). They instructed preschool children in day schools to complete tasks at two different levels of difficulty. Conditions of reinforcement and no reinforcement were then presented. It was found that compliance with easy tasks was maintained without reinforcement but that compliance with complex tasks remained high only when the specific task was reinforced. Bucher and Okovita conclude that the degree of attention which a child pays to the teacher's instructions varies with task difficulty under natural conditions.

Findings of this study suggest that it is possible to have a student persevere on tasks with materials inappropriately matched to his ability, provided he is reinforced for doing so. Punishment may also keep a student on task. However, in either case, comprehension and achievement would, to varying degrees, be sacrificed.

The importance of stimulus complexity in holding the attention of a child was stressed by Moyer and Gilmer (1954). They suggest that the skill level of a particular child is a very important variable influencing attention span, especially in such word tasks as reading and writing. It appears that the amount of time a child attends to a task is affected by the ability of the child to perform the task. The more complex the task, the less able the child, the shorter the time spent attending. Moyer and Gilmer reach the conclusion that attention spans may be increased by providing the right task for the particular skill and age level of the child and remark as follows:

It may well be that a word task which is made sufficiently interesting to the child can, within his range of abilities, evoke attention spans well beyond the limits generally accepted. (p. 466)

Interrelationships Among Readability, Achievement, and Behavior

A specific form of task difficulty is readability, or the relative difficulty level of a written passage. Readability is influenced by such factors as sentence length, concept difficulty, and semantics, all of which contribute to a passage's complexity. To date, most

research has dealt with readability as it affects academic achievement. Other studies relate academic achievement to classroom behavior. Both of these fields of study have made valuable contributions to the understanding of readability, achievement, and behavior.

In 1943, Jones sought to find out if adapting a curriculum to the individual differences of the pupils would have any effect on their progress in school. Two groups of fourth graders were formed. Assessment of academic ability was conducted with the experimental group and materials of corresponding difficulty level were assigned. The control group used only materials labeled "fourth grade." Jones found that increased achievement in reading for superior, normal, and dull children resulted from matching difficulty level and nature of the material to the ability and interest of the individual student.

Recent studies into the relationship between readability and achievement have confirmed Jones' findings. Ewing (1976) asked 169 ninth grade students from two rural counties in north Florida to read either of two state driver's handbooks and prepare for a comprehensive test on what they learned. One of the handbooks was written at approximately fifth grade readability level while the other was approximately eleventh grade. Ewing found that students working from the handbook with lower readability had higher mean scores on the comprehension test than students using the higher readability handbook. He concluded that a student's level of reading ability was a critical factor affecting achievement on the exams.

Final course grade was the dependent variable used by Bertalan (1977) to determine the effect of readability on the achievement of community college students. He found that significantly lower final course grades were obtained by students who possessed reading grade placement levels below the readability levels of their assigned textbook. He also found that a student's final course grade increased as his reading grade level increased over the readability level of the materials being used.

Bankston (1975) obtained a similar result using pre- and post-tests of mathematics achievement. Community college students who had received mathematics materials written at the lowest readability level (least difficult) scored significantly higher on the posttest than a group with materials at a mid-difficulty level and still higher than a group with materials at the highest level of readability.

If lower achievement results from a student's use of levels of difficulty exceeding his reading ability, then can reading comprehension and academic achievement be increased by lowering the level of readability of a textbook? Paige (1978) conducted a study in which two readings from an industrial arts electronics text were rewritten at approximately an eighth grade level of readability. Both the original readings and the rewritten passages were then randomly assigned to tenth, eleventh, and twelfth grade students who were later tested. Paige found that lowering the level of readability substantially increased the level of comprehension for these students.

The above mentioned studies conclude that readability is a significant variable affecting achievement. Achievement, particularly when it is lacking, has in turn a strong effect on behavior. Many forms of deficits in academic achievement, and especially reading disabilities, have long been recognized as significant factors contributing to inappropriate classroom behavior and other forms of personal maladjustment (Fendrick & Bond, 1936; Gates, 1933; Robeck & Wilson, 1974; Tinker & McCullough, 1975). One estimate (Gates, 1947) placed 75 percent of severe cases of reading disability in the category of expected to exhibit personality maladjustment.

According to Bond and Tinker (1973), repeated failure in attempting to obtain adequate reading skills and the resulting frustration which occurs is a primary cause of emotional maladjustment. This frustration leads to aggressive anti-social behavior as well as shyness and daydreaming. In effect, a pattern develops in which the inability to read adequately leads to frustration with school which in turn leads to dropping out, reduced employment opportunities, and possible delinquency.

The long-range consequences of failure in reading, e.g., unemployment, crime, etc., are severe and usually debilitating. It is, however, the more immediate affect of reading disability within the classroom that is of principle interest to this study. These affects manifest themselves in easily observed behaviors as well as feelings and attitudes which are often obtained through self-report. Several longitudinal studies have brought these behaviors and feelings to light.

A four-year study by Jessor and Jessor (1977) revealed that the student who does not value academic achievement and who does not expect to do well academically is more likely to engage in problem behaviors. Feldhusen, Thurston, and Benning (1973) conducted a longitudinal study involving 1,550 children in grades three, six, and nine. School factors were found to be significantly related to aggressive problem behavior. Children identified as behavior problems had significantly lower reading and mathematics achievement scores and were significantly more inclined to drop out of school.

Other studies confirm the close association between academic achievement deficits and conduct problems, carefully avoiding, however, any conclusions as to the necessary precursor. Camp and Zimet (1975) observed the behavior of high, middle, and low reading group first graders during their reading periods and found that as reading skill decreased, deviant behavior and classroom interruptions increased. Glavin and Annesley (1971) found that 50 percent of the 130 behavior problem boys in their sample were extreme under-achievers in reading as compared to 21 percent of the normal boys. Further studies have shown that unsuccessful readers exhibit a lack of social confidence and low motivation toward school achievement (Karlsen, 1955), are poorly adjusted to rules (Granzow, 1954), and are more unable to acknowledge or accept blame than successful readers (Spache, 1957).

It becomes evident from the studies presented in this section that there exists a school-based chain of events in which a student can become entwined. Readability levels that significantly exceed

an individual's ability create achievement deficits, especially reading disabilities. These, in turn, contribute to classroom behavior problems. However, even though this association of events has long been recognized, the research literature remains noticeably lacking in studies concerning the direct effect of readability on classroom behavior.

Teacher Practices Affecting Classroom Behavior

One of the most often heard calls for help coming from today's public school classrooms is an impassioned plea for assistance with behaviorally disordered children. According to a National Education Association survey (1976), the number one problem facing teachers is discipline in the classroom. An earlier survey of teachers by Gilmore and Chandy (1973) found that teachers agree on the type of child who most needs psychological referral--the behavior problem.

The reported incidence of behavior problems in the classroom is high. Basing their findings on years of observing children, Swift and Spivack (1975) suggest that from 25 to 30 percent of the children in a typical classroom display some kind of problem behavior. Inattentiveness, apathy, and withdrawal from class activities were most frequently listed.

An important link exists between teacher practice and student behavior problems. Inability to recognize readiness levels, promoting children by age rather than achievement, attempting to adjust the child to the curriculum rather than the curriculum to the child, and failure to provide meaningful and enjoyable academic materials

are just a few common teacher practices which affect learning success and classroom behavior. Many other daily educational practices require continuing evaluation to determine their current suitability for individual students. According to Stradley and Aspinall (1975), a direct relationship exists between classroom functioning demands, potential student frustrations, and consequent discipline problems. Improvement in classroom behavior can be realized if the teacher devotes a sufficient amount of time to the selection of classroom learning activities most likely to insure success for each individual student.

A logical first place for teachers to look for an explanation of problem behavior is within their own classroom (Skinner, 1968). A teacher's workday consists of the direct application of educational and behavioral principles to children's behavior. It is these practices which should be examined. Speculating on original causation is not as educationally relevant as concentrating on finding what factors are currently operating that interfere with improvement in a child's learning or behavior.

The need for teachers to look closely at their own teaching practices finds particular relevance when considering attribution of causation. Baldwin, Johnson, and Wiley (1970) asked teachers why they thought students succeeded or failed in school. They learned that teachers attribute pupil success in certain subject matters to variables such as methodology, which are controlled by and the responsibility of the teacher. On the other hand, when students fail in certain subjects, they see the reason as being due to the individual students. That is, the problem is intra-individual

and involves deficits in intelligence, readiness, motivation, or attention. Baldwin et al. conclude that teachers tend to accept credit for success in their pupils but attribute failure to causes over which they have no control.

Teacher practices have been scrutinized and empirical evidence of their influence obtained. A study was conducted by Chall and Feldman (1966) in which observations of teacher practices during reading lessons were made over a period of 140 days of classroom instruction. They were concerned with the affect of certain teacher-controlled variables on reading achievement. First grade children from socially disadvantaged neighborhoods in New York City were administered a total of 45 measures of reading readiness and achievement during the school year. An analysis of variance showed a significant association between appropriateness of difficulty level for the reading lessons and pupil reading achievement. Chall and Feldman concluded that one of the most important teacher practices affecting reading achievement in the first grade was accurately matching lesson difficulty to student ability.

Other studies show that if a teacher increases the amount of attention she pays to a student, the attending behavior of that student can be increased (Brodén, Bruce, Mitchell, Carter, & Hall, 1970). Furthermore, reinforcement of academic performance by the teacher can reduce the rate of disruptive behaviors occurring in the classroom (Ayllon & Roberts, 1974).

It is possible that in many instances teachers are unaware of their part in fostering student behavior problems. When this is the case, children are inadvertently subjected to practices having a

negative effect on their behavior. The practice of mismatching student ability and material difficulty is a common occurrence and one which can easily remain unrealized.

Readability of Academic Materials

The readability of textbooks has long been of concern to many psychologists and educators. Notable among these is E. L. Thorndike, whose research studies have greatly influenced the design of educational books and other reading materials. Thorndike was interested in word difficulty and its effect on the ease with which material could be read. By the 1920's, he had inventoried a large sample of words and determined the frequency with which they occurred in print. This frequency count was used as a measure of the difficulty of the word. Thorndike proceeded to modify a great number of literary works, many of them classics, in order to make them readable for children.

In the 1930's, a student of Thorndike's, A. I. Gates, began to design readers for the lower grades. Readability designations were based on Thorndike's notion of the frequency with which vocabulary words occurred. Each book which Gates designed had words occurring with sufficient frequency to insure a degree of mastery by the children.

Since the 1930's, interest in readability and reading materials has steadily grown. It was soon realized that frequency of vocabulary words was not a sufficient condition for determining readability.

Other factors such as sentence length and syntax were also important variables. Readability formulae were developed to account for this (Klare, 1963). These formulae differentially weigh varying factors considered to be instrumental to determining the difficulty of the material. A grade level score, usually based on sample passages from the text, is derived and represents the readability of the full text.

The basic tool of the teacher, the textbook, continues to deserve attention as a subject for evaluation and research. However, many teachers appear to be unaware of the need to evaluate the reading materials they use daily, especially with regard to readability. Eberwein (1979) reported the results of a study in which he found that the average elementary school reading text, which was designated at a particular level of difficulty, actually had a range of 4.65 levels. When teachers who use the textbooks were given information on the readability of these books, they were surprised with the variations, found the variations more than expected, and were concerned about the instructional implications.

Jorgenson (1975) also found that "elementary school teachers vary widely in their ability to judge accurately the difficulty level of paragraphs from various grades, and that a common sense of 'grade level' does not exist" (p. 70).

Textbook readability becomes an especially critical factor affecting learning success and classroom behavior when considering that the range of reading abilities in any one classroom is great and becomes greater each successive grade (Goodlad, 1966). The spread of high to low reading achievement can range from one and

one-half, to twice the number of the grade level. A third grade classroom, therefore, frequently contains students working at first grade level as well as students at a sixth grade level.

This continually widening gap between poor readers and good readers in a classroom presents a significant instructional problem for the teacher. The most common approach to reading instruction, used in an estimated 80 to 95 percent of American schools, is the basal reader method (Chall, 1967; Staiger, 1969). This approach generally involves dividing a class into three reading groups based on ability and assigning either a low, middle, or high level reading text. Three levels, however, cannot account for the wide range of reading ability existing within the typical classroom. A portion of the class will therefore be inappropriately suited to the reading materials assigned to them.

Inaccuracies of Grade Level Designations in Basal Readers

A teacher can become aware of the wide range of ability within her classroom after having individually evaluated each student's reading ability. She can then decide to preclude using the standard three reading group system and instead opt for individualized instruction with reading texts whose grade level designations are "appropriate" for each child. The teacher would suppose that this method would produce an accurate match of student ability and material difficulty. Unfortunately, it often does not. The reason: the publisher has failed to accurately assess the readability of the text and so the grade level designation is incorrect. In a study in which 200 books were graded for readability, Mills and Richardson

(1963) found that approximately half of the books were incorrectly labeled as to grade level. Two formulae which produced comparable results were used to determine readability in the study. In most cases, the publishers underestimated the difficulty level of the books from one to four grades. Mills and Richardson also found that about one half of the publishers used no standardized formulae but rather relied on the judgment of their authors or educational consultants to determine the grade level of their texts.

A computerized readability assessment program was devised by Britton and Lumpkin (1977). Its purpose was to verify the publisher designated reading levels of approximately 200 stories in sixth grade reading texts. Four readability formulae were incorporated into one program enabling comparative analysis. Among the findings of this study were that sixth grade texts produced by the five major children's textbook publishers studied (Ginn, Economy, Houghton Mifflin, Merrill, and Scott Foresman) had reading grade levels ranging from 3.0 to 16.0, that as much as 63 percent of the stories in one of the readers were two or more grade levels above the publisher designated grade level, and that stories in the five texts were not arranged in order of difficulty but rather some of the most difficult stories were found in the beginning of the book and some of the easiest ones found toward the end. Britton and Lumpkin conclude:

Textbook products that shape attitudes, self-concepts, and foster the success or failure of our teachers and children need to share the same labeling and product assessment standards as do tires or a bottle of aspirin. (p. 196)

Further evidence exists which indicates that basal reader's grade designations do not accurately reflect their difficulty. Bradley and Ames (1977) cite studies which have investigated the many factors contributing to a book's readability and conclude that publishers had failed to reliably account for them. Pruitt (1977) found the average range of readability within fourth, fifth, and sixth grade social studies texts to be 6.2 years.

The research thus indicates that in most cases basal readers are more difficult than the grade level assigned to them by the publisher. Assigned grade levels are therefore underestimates of their readability. Without realizing it, teachers are assigning their students reading tasks which are inappropriate and unreasonable. The continual stress from this practice could be responsible for a sizable amount of the classroom behavior problems observed daily.

Mismatch of Academic Task and Student Ability

There is general agreement within educational circles that a need exists to adapt the school curriculum to individual student requirements. It is assumed that the provision of academic materials commensurate with ability is a necessary ingredient in the recipe for student learning. However, research indicates that a discrepancy exists between the assumption and actual practice. A lack of agreement between student ability level and the difficulty level of academic materials is a common occurrence in public schools. Jorgenson (1977) and Jorgenson, Klein, and Kumar (1977) found wide discrepancies between the abilities of elementary school students

and the readability of materials to which they were assigned. Some students were assigned materials that were much too difficult for them while others were working with very easy materials. It was apparent that elementary school teachers did not consistently match materials to ability. Jorgenson concludes from his studies that "teachers tend to fit students to the level of material, rather than selecting materials to match student ability" (1978). In other words, it is the grade level of the student to which reading materials are matched rather than individual ability level.

Jorgenson's conclusions are borne out by the findings of other studies. In Pruitt's (1977) sample of 1,467 elementary school students, 58 percent were found to be assigned to textbooks at a level in which minimum comprehension and a high degree of frustration occurred. Campbell (1973) studied 455 fourth, fifth, and sixth grade students and the textbooks assigned to them. One-fifth of the students were found to be unable to comprehend their social studies texts and over one-half of the fourth grade and one-fourth of the fifth and sixth grades could not adequately comprehend their science texts.

In addition to finding that most of the freshmen college textbooks he investigated were written on a level higher than freshman students were capable of reading independently, Fox (1978) also found that most instructors were unaware of the types of problems, both learning and emotional, that students incurred as a result of the mismatch.

The ability or inability of teachers to judge the difficulty level of reading materials was further researched by Miller and

Marshall (1978). They found that teachers frequently have difficulty making an accurate match between student ability and the readability of written materials. They see the problem as being due in large measure to teacher training programs which are inadequate in this area and need to be changed.

Finally, two studies, one involving non-Anglo-Americans, and the other disabled readers, point out that for special populations, student ability-material difficulty matching can be an especially critical variable. Auvenshine (1978) compared the reading abilities of three groups of junior college students with their textbooks. He found the mean readability levels of the texts to be one to two levels above the reading ability of the Anglo-American sample, three to four levels above the Mexican-American sample, and four to five levels above the Black-American sample. Apparently, despite the fact that all of the students were sufficiently qualified for admittance to junior college, a large discrepancy in reading ability exists between various ethnic groups. When the extent of the mismatch is realized, it becomes easier to understand the confusion and frustration minority students must encounter in their daily academic schedules.

While being assigned exceedingly difficult reading tasks is a major problem for the average reader, it is an even greater aggravation for the already disabled reader. Arnold and Sherry (1975) compared reading levels of disabled third, fourth, and fifth grade readers with their assigned textbooks and found sizable discrepancies between ability level and task difficulty. They conclude that

disabled readers are placed in books the difficulty of which is bound to frustrate them and that this is an "indefensible educational practice."

This concludes the literature on the concept of ability/difficulty matching. The following section deals with the practical classroom implications of mismatching.

Relationship of Ability-Difficulty Matching to Classroom Behavior

Only three studies dealing with the relationship between student ability, task difficulty, and classroom behavior are known to this researcher. Dr. Gerald Jorgenson of John Carroll University is responsible for all three of these. In his first study (1977), Jorgenson determined the reading ability of 71 second through sixth grade students by use of two standardized reading tests. He then used readability formulae to determine the difficulty level of reading material currently in use by the students. A measure of classroom behavior was obtained by having teachers complete a behavior rating scale for each student in their classes who participated in the study. Jorgenson then obtained a measure of the accuracy of the match between ability level and material difficulty level by subtracting the reading test grade level score from the grade level score yielded by the readability formula. Thus, a "difference score" was available for comparison with the classroom adjustment scores obtained from the behavior rating scale.

Jorgenson found that as the reading material became more difficult (as the difference scores increased) the students were judged by their teachers as becoming more disturbing to the classroom, impatient, and reliant for directions on persons other than themselves. As such, student behavior was judged less appropriate as material became more difficult for students. Conversely, as material became easier for students, teachers tended to rate student behavior as more appropriate.

A second important finding by Jorgenson was that teachers were not consistently using material matched to the ability levels of their students. Rather, they were assigning material roughly according to the grade level of the students.

A second study (Jorgenson, Klein, and Kumar, 1977) and a third (Jorgenson, 1978) were essentially replications of the first using different measuring instruments and student samples. Like the first, the procedure involved determining student ability through the use of a standardized reading or achievement test, obtaining a measure of readability (the second study used only the child's assigned reader as his level of instruction; the third, a readability graph), then having the teachers involved complete a behavior rating scale, and finally comparing difference scores to behavior ratings. Both follow-up studies confirmed the original findings, thus extending their generalizability.

The Jorgenson studies all share several methodological weaknesses affecting the conclusiveness of their findings. One of these weaknesses is the use of standardized reading and achievement

tests to determine student ability. These tests are often only indirectly related to the actual performance of pupils in a classroom and frequently use reading passages which are dissimilar to the type of reading material available for instruction in the classroom (May, 1973). Furthermore, grade-equivalent scores on these tests are often overly high estimates of a child's ability and in some cases are quite severe in their normative standards for beginning readers (Harris & Sipay, 1975). In addition, since the test scores are infrequently gathered (there is often only one test administration), they can easily misrepresent true ability (Lovitt & Hansen, 1976). Having only one or two data points permits confounding by such variables as environmental, physical, and emotional conditions on test day, maturation, and history.

A second weakness is the use of difference scores. According to Thorndike and Hagen (1969), difference scores are considerably less reliable than single scores. The errors of measurement on each test contribute to error variance in the difference scores and the true variance that the two tests measure in common reduces the variability of the difference scores.

The use of readability formulae to determine the difficulty level of reading materials presents another problem. As mentioned earlier, different formulae often produce different grade level scores for the same reading passage (Bormuth, 1966). More importantly, however, is that material is evaluated in terms of an averaged standard of performance rather than in comparison to the ability of a particular individual. Averaging masks individual

differences. Material at a particular grade level designation may therefore be too difficult for one reader and too easy for a second, even though both readers were assessed to be at the same level of ability.

Finally, behavior rating scales, such as those used by Jorgenson, are beset with problems of reliability and validity. Symonds (1931) summarized a number of studies which show that the between-rater reliability of conventional rating procedures is low and concluded that rating scales were generally inconsistent measures of behavior. Thorndike and Hagen (1969) suggest that this conclusion is still warranted after the lapse of years and go on to say that validity becomes questionable when rater unreliability and biases are considered.

Among the many problems involved in obtaining a sound measure of behavior from rating scales are: (1) unwillingness on the part of the rater to take the necessary pains in completing a careful and thoughtful rating, (2) identification of the rater with the person being rated, (3) limited opportunity to observe the person rated in all the various areas called for on a scale, (4) covertness of some traits that are measured, (5) ambiguous definitions of dimensions being measured (one scale used by Jorgenson has a factor called Impatience, which is defined as "concerned with an inappropriate drive to enter into and complete the work assigned"), and (6) specific rater idiosyncracies including deficits in memory.

These deficiencies in rating scales result in errors which occur so frequently and persistently that special efforts are needed to counteract their influence (Gronlund, 1972). However,

special efforts are rarely employed and errors of personal bias, response set, and halo effect are among the many that result.

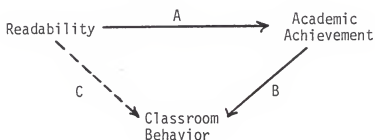
A viable alternative is available, however, which corrects for most, if not all, of the weaknesses of behavior rating scales. This alternative is direct measurement. Cobb (1972) states that rather than concentrate on possible internal mediating variables inferred by teacher's ratings, emphasis can be placed on the frequent observation and recording of the child's overt classroom behavior. In this way it is actual behavior which is the dependent variable rather than a scale rating requiring an inference to behavior.

Other advantages of direct measurement include the use of clear operational definitions rather than complex hypothetical constructs and frequent, continuous data collections which provide a picture of behavior over time. Findings from direct measurement are directly applicable to the individual(s) involved and as such are of practical use to the classroom teacher.

Summary of Review of Literature

This review has examined several factors with varying potential to effect the relationship between student ability/material difficulty matching and classroom behavior. The studies interrelating task difficulty, achievement, and behavior are particularly noteworthy because they point out that research efforts have usually been only tangential to the direct relationship between task difficulty and

behavior. That is, the great majority of investigative effort has been directed toward either the relationship between task difficulty and achievement or between achievement and behavior. The link between task difficulty and behavior has itself received a paucity of research attention. The diagram below shows the interrelationship of variables with regard to research emphasis. Relationship C is the major concern of this study.



The literature, therefore, while providing strong indication that task difficulty can adversely affect achievement and that poor achievement often adversely affects classroom behavior, still only provides indirect evidence of a relationship between task difficulty and behavior.

The review next discussed the importance of teacher practices as a factor affecting achievement and behavior. Special emphasis was given to the notion that it is very easy for the teacher to overlook the influence of her own varying educational practices and thus inadvertently contribute to classroom problems.

The form of task difficulty most relevant to this study is readability, and particularly the readability of basal readers. This review presented findings which showed that publisher designated grade levels of basal readers were frequently inaccurate. Most

teachers accept publisher grade level designations as correct. Assignments to readers of inappropriate difficulty levels are therefore frequently made. Abundant evidence of this mismatch of task difficulty and student ability exists and a portion of it was presented.

Few researchers have attempted to quantify the amount of "frustration" experienced when a mismatch exists and it remains a subject for speculation. To this writer's knowledge, none have observed and counted behaviors possibly indicative of frustration in such situations of mismatch.

The only studies dealing directly with the issue of matching and classroom behavior have been those of Jorgenson (1977, 1978) and Jorgenson et al. (1977). The findings of these studies were shown to have limited conclusiveness due to the methods chosen for data collection and interpretation. A need exists, therefore, to further explore the issue using alternative research methods. This study seeks to respond to that need.

CHAPTER III

METHOD AND PROCEDURES

Subjects

Five third grade children attending public school in Alachua County, Florida, participated as subjects in this study. These children were randomly drawn from middle and lower reading groups that had been established within two classrooms. A description of the subjects with regard to age, sex, IQ, ethnic origin, and reading group placement can be found in Appendix B.

Setting

All experimental conditions took place within the regular classroom setting. Children sat in their assigned seats and conditions were implemented during morning sessions regularly scheduled for reading and other seatwork.

Materials

Reading passages were selected from the Ginn Reading Series (Clymer & others, 1976) and Beyond Counseling and Therapy (Carkhuff & Berenson,

1977). The Ginn Series is a basal reading series containing 15 books (Levels 1-15). Each book purports to represent a higher level of reading difficulty than its predecessor in the series. For example, Level 8 represents the grade equivalent of the first half of third grade, Level 9 the second half of third grade, Level 10 corresponds to fourth grade, Level 11 to fifth grade, etc. However, for the purposes of this study, all grade level designations were ignored. The books were used to supply differing reading passages with which a functional assessment of material difficulty and student ability was made. The Ginn is the series currently in use by Alachua County elementary schools. The Carkhuff book was required to provide a high level of difficulty unavailable in the Ginn series.

Dependent Variable

Directly observed off-task behavior (see Appendix A) is the dependent variable in this study. The unit of measurement for the dependent variable is duration, i.e., the number of seconds off-task during 5 and 7 1/2 minute reading samples obtained over successive school days.

In addition to off-task behavior, an estimate of interest in or satisfaction with the material being read was sought. This was obtained by asking each child to rate each passage immediately after reading it. Responses were classified according to a four choice scale.

Assessment of Student Ability and
Level of Material Difficulty

This study assessed student ability and material difficulty by use of an idiographic rather than a nomothetic approach. Children were not compared to other children to obtain a measure of their ability, nor was material compared to some normative standard for a measure of its difficulty. Accordingly, no standardized achievement or aptitude tests or readability formulae were used. Instead, each student's reading ability was assessed as a function of his proficiency with the specific reading material he was using. In this way also, the difficulty of specific reading material was determined by assessing the degree with which an individual could read the material.

The procedure used for determining student ability and reading material difficulty was the Reading Placement Inventory (RPI), developed by Lovitt and Hansen in 1976, and modified slightly in 1978. The RPI consists of the following six steps:

1. Select a reading series whose books are designated as graduated in ability.
2. Excerpt three 100-word samples of text, one each from the beginning, middle, and end of each reader.
3. Compose four comprehension questions (one recall, one sequence, one interpretation, and one vocabulary) for each 100-word passage.
4. Collect information on the child's reading ability through individual testing over a period of days. Each child is asked to orally read a passage and

answer the four comprehension questions accompanying it. A record is kept of each misread word. Incorrect words include omissions, substitutions, insertions, mispronunciations, hesitations longer than four seconds, and "I don't know." Proper names and places, repetitions, and self-corrections are not counted as incorrect responses.

5. Calculate the mean correct words per minute (CWPM), incorrect words per minute (IWPM), and comprehension percentage (Comp %) for the three 100-word passages using the following formulae:

$$\text{CWPM} = \frac{300 - \text{number of incorrectly read words}}{\text{Total number of seconds read}} \times 60$$

$$\text{IWPM} = \frac{\text{Total number of incorrectly read words}}{\text{Total number of seconds read}} \times 60$$

$$\text{Comp \%} = \frac{\text{Number of correct answers}}{\text{Total number of questions}}$$

6. Placement and subsequent future confirmations.

Although student ability and material difficulty was assessed orally and subjects then asked to read silently during experimental conditions, research has demonstrated a rather high relationship between silent reading comprehension and accuracy of word recognition in oral reading. Bond and Dykstra (1967), for example, found correlations of .78 to .85.

Independent Variable

Level of reading material difficulty as determined by the RPI for each subject was the independent variable in this study. Each child was experimentally subjected to three ranges of the independent variable. The criteria for each of the three difficulty levels are as follows:

- Level 1: eighty and over CWPM
and/or
three or fewer IWPM
76 percent or better Comp %
- Level 2: forty-five to seventy-nine CWPM
and/or
four to eight IWPM
50 to 75 percent Comp %
- Level 3: twenty to forty-four CWPM
and/or
nine or more IWPM
0 to 49 percent Comp %

For the purpose of this study, Level 1 was considered the difficulty level at which a child can read fluently and with high comprehension. As such, it was the baseline condition for silent reading performance. Criteria for Level 1 are roughly comparable to a 30 percent improvement over criteria for Level 2. Criteria for Level 2 were determined empirically by Lovitt and Hansen (1976) to be appropriate for reading performance with instructional assistance. Level 3 represents an additional proportionate decline in reading performance and thusly an increase in material difficulty.

Experimental Design

The purpose of the design used in this study was to allow one to draw convincing conclusions concerning the effects of difficulty level on off-task behaviors. A minimum of three Ginn readers with functionally different difficulty levels (as determined by the RPI) were used with each subject. During each day of experimental conditions (school days) the subject read silently for five minutes in each of the differing readers. The five minute reading periods occurred in succession with a one minute break between reading sessions. A second phase consisted of two levels read for 7 1/2 minutes each.

Of the three or more levels of difficulty used, Level 1 was called the baseline level. While it was referred to as baseline it should not be considered a separate experimental condition. It is, rather, one of the values of the independent variable. Since it is a level at which a minimum of off-task behavior is presumed to occur, it will be used for comparison with the two more difficult levels.

A multi-element within sessions design (Sidman, 1960) was selected for use in this study. According to Bailey (1977) "the design operates primarily by bringing the same behavior (of one or several subjects) under the control of several different experimental procedures. . . . When shown graphically, the data points for each of the separate conditions are connected and any systematic deviation from the baseline demonstrates experimental control" (p. 162).

In contrast to most single case designs, the multi-element design contains a continuing current baseline to which treatment effects can be compared. For example, instead of having ten days of baseline conditions and ten days of treatment conditions, the multi-element format would contain baseline and treatment conditions occurring each day for 20 days. The effects of different values of the independent variable can therefore be compared on a daily basis. This feature permits control of temporal variations which contribute to intrasubject variability. According to Sidman (1960), "the repeated replication of each element of the baseline permits an otherwise impossible degree of confidence in the adequacy of the experimental control" (p. 325). This confidence stems in part from the fact that each succeeding day can act as a replication of the experimental effect. The design is portrayed schematically in Figure 1.

<u>Subjects</u>	<u>Order of Treatment</u>					
1	L ₁ L ₂ L ₃	L ₃ L ₁ L ₂	L ₂ L ₁ L ₃ *
2	L ₁ L ₂ L ₃	L ₃ L ₁ L ₂	L ₂ L ₁ L ₃
3	L ₁ L ₂ L ₃	L ₃ L ₁ L ₂	L ₂ L ₁ L ₃
4	L ₁ L ₂ L ₃	L ₃ L ₁ L ₂	L ₂ L ₁ L ₃
5	L ₁ L ₂ L ₃	L ₃ L ₁ L ₂	L ₂ L ₁ L ₃
	Day 1	Day 2	Day 3	Day n

Key:

L₁ = reading difficulty level 1

L₂ = reading difficulty level 2

L₃ = reading difficulty level 3

*Order of treatment levels continues to be randomized.

Figure 1

Schematic Illustration of a Multi-Element
Within Sessions Design

Experimental sessions were continued over a period of time sufficient to achieve stability in the data. Sidman (1960) defines a stable, or steady, state as one in which the behavior in question does not change its characteristics over a period of time. Bailey (1977) states that since a visual inspection of the data is the primary basis for drawing conclusions, any condition should run until one can begin to predict where the next data point will be.

There are several reasons why this design was selected for use in this study:

1. It is well suited to the analysis of several values of the independent variable since sequence and order effects are corrected for by the random presentation of conditions (Bailey, 1977).
2. It yields frequent samples of behavior and allows for repeated replication of treatment effects across days and subjects (Sidman, 1960).
3. Interaction effects from ongoing environmental changes, e.g., maturation, history, etc., are minimized by repeated daily exposure of each subject to all experimental conditions (Sidman, 1960).
4. Since frequent and repeated time samples of each element are provided, any loss of experimental control that may occur becomes immediately evident and can be taken into account when evaluating the data (Sidman, 1960).

Observers

The data collectors were students in the Department of Special Education of the College of Education, University of Florida. They were trained as behavioral observers by the researcher in pre-experimental training sessions. Training consisted of complete familiarization with (1) the behavioral definitions of the dependent variable, (2) the process of observing the dependent variable, (3) the recording form, and (4) the stopwatch. Instruction in polite and unobtrusive classroom behaviors was also included.

The pre-experimental training period took place over several days. During this time, the researcher and each of the observers conducted simultaneous observations of selected children. Data records were compared, dissimilarities discussed, and necessary corrections made.

In this study one observer (A) observed two subjects and the second observer (B) observed three subjects. During actual experimental conditions the author carried out reliability checks on 25 percent of each observer's total number of observation sessions. According to Bailey (1977), a reliability check is performed when two people observe the same behavior simultaneously, using the same definition and scoring system. Observer reliability was obtained by the following formula for duration measures (Bailey, 1977):

$$\text{Percent of Agreement} = \frac{\text{smaller score}}{\text{larger score}} \times 100$$

A 90 percent agreement between observer and researcher for both pre-experimental and experimental conditions was sought.

Data Collection

Data were collected during daily observations of the children in their regular classroom environment. Collection procedures were as follows: the observer presented the subject with a particular level of reading material and returned to the observation point. A stopwatch was started immediately upon return and allowed to run without interruption for five minutes. An occasional glance at this watch permitted the observer to keep track of time elapsed in the session. A second stopwatch was started whenever off-task behavior was observed. Total number of seconds recorded on this second watch constitutes the dependent measure. This recording method provides a maximum amount of continual observation time with a minimum amount of recording distraction.

Duration of off-task behavior, completion time, ratings, and comments were recorded immediately after the five minute session had elapsed. A form devised for the purposes of this study was used to record the data.

In order to avoid possible experimental bias, neither the experimental questions nor the purpose of the study was discussed with the observers until after all conditions had been completed.

Data Analysis and Interpretation

Data recorded during observation periods are presented graphically in clear form in the next chapter. Data underwent median smoothing procedures in order to present a more easily understood curve of performance. The procedure, as described by Tukey (1977), involves obtaining medians of three-day data intervals and plotting these points. The smoothed data eliminates erratic fluctuations and facilitates recognition of tendencies in the data.

CHAPTER IV

RESULTS

This study sought to determine if a relationship exists between two variables--academic task difficulty and behaviors defined as off-task for an elementary school classroom. Total occurrence of off-task behavior was determined by directly observing five children from two third grade classrooms. A duration measure was the basic datum used, that is, the amount of time that the individual emits a given behavior was the datum for "off-task" behavior.

In addition to the principal duration data, a measure of interest in, or satisfaction with, the material read in experimental sessions was sought. This information was obtained by simply asking the child how he or she liked each of the reading passages and recording these answers. Answers were then classified according to four possible scaled responses.

After approximately 8 to 10 days of experimental conditions, it became evident that two subjects (III and IV) maintained almost total perseverance to task regardless of difficulty level. In addition, for two of the subjects (I and IV), the functional difference between difficulty levels for texts at the upper ranges of the Ginn Series was minimal. That is, while reader levels 13 and 14 met criteria for difficulty levels 2 and 3 respectively, the number of

correct words per minute, incorrect words per minute, and comprehension percentage clustered around the cut off criteria separating the two levels. For these reasons, an additional experimental phase was added to the study.

In Phase II, two 7 1/2 minute sessions were presented to all but one subject. The one child excluded, Child II, was not using the higher level Ginn texts and had already provided data indicating a clear experimental effect. Of these two 7 1/2 minute sessions, one was a duplication of the Level 1 task extended 2 1/2 minutes. The other session involved presenting the subject with an exceptionally difficult reading task, one in which the child was assessed to read 20 or fewer correct words per minute with 25 percent or less comprehension. To obtain this level of difficulty for Subjects I and IV, a complex college level text was used. For the remaining two children reading at considerably lower levels of difficulty, level 14 of the Ginn series sufficed.

Reliability Measure

The method used to determine reliability for duration measures (Bailey, 1977) was described in the previous chapter. Initial reliabilities with one subject (Child II) failed to meet criterion and necessitated additional clarification of the dependent variable with Observer A. Total accumulated checks for reliability of observations between researcher and observers resulted in a 90 percent agreement for Observer A and a 93 percent agreement for Observer B. Checks in which both the observer and the checker obtained duration measures of less than ten seconds were discarded. This procedure eliminated the exorbitant influence on reliability percentages which results

from small time differences having no practical significance. While durations of less than 10 seconds were excluded when computing reliability percentages, they were included on all occasions as experimental data points. The mean duration of off-task behaviors observed during reliability checks, excluding those under 10 seconds, was 70.4 seconds.

Effects of Task Difficulty Level on Off-Task Behavior

Results of this investigation are displayed graphically in Figures 2 through 4. In order to even out the erratic highs and lows and present a more easily understood curve of performance, the data are presented after undergoing a three-day median smoothing (Tukey, 1977). Smoothing is defined as the process in which a sequence of smoother appearing values is calculated from the given values. In this case, the medians of days 1, 2, 3; 2, 3, 4; 3, 4, 5; and so on are derived and plotted. Midpoints of the first two days and last two days represent first and last plotted points respectively.

Essentially no off-task seconds were recorded for Subjects IV and V. The reader can easily visualize the flat-line graphs for these two subjects and therefore they are not offered. Table 1 presents the mean number of seconds and mean percentage of off-task behavior for all five subjects.

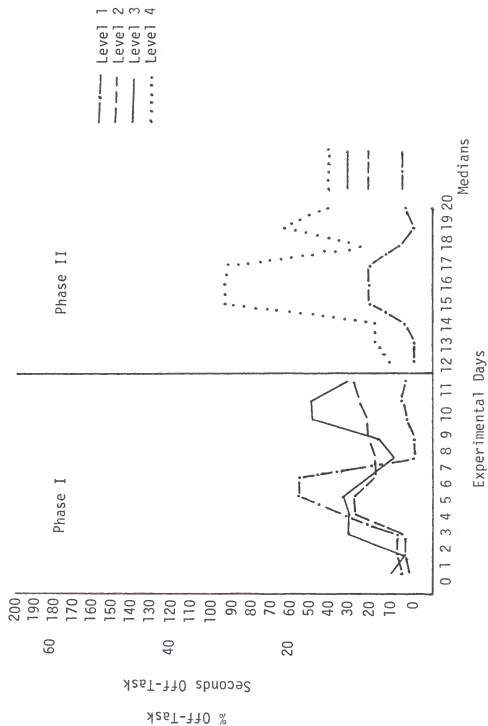


Figure 2
The Effect of Varying Levels of Task
Difficulty on Duration of Off-Task Behavior

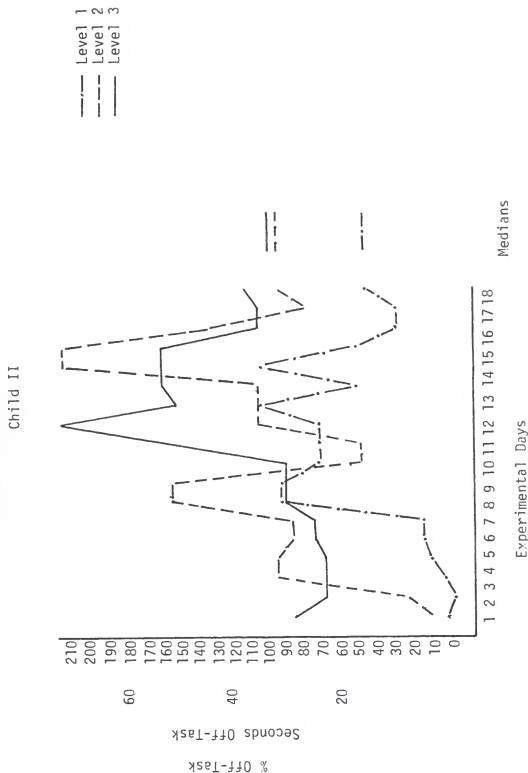


Figure 3

The Effect of Varying Levels of Task
Difficulty on Duration of Off-Task Behavior

Child III

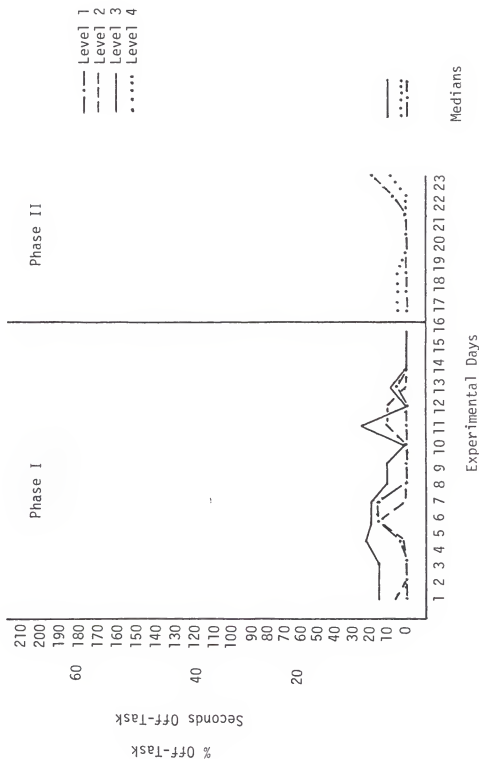


Figure 4
The Effect of Varying Levels of Task
Difficulty on Duration of Off-Task Behavior

Table 1
Mean Percentage and Mean Number of Seconds
of Off-Task Behavior by Level and Phase

Subject		Level				
		1	Phase I 2	3	Phase II 1	4
I	seconds	23.7	21.1	27.3	9.7	51.4
	percent	.079	.070	.091	.021	.114
II	seconds	60.9	98.3	121.6	-	-
	percent	.203	.328	.405	-	-
III	seconds	6.7	5.8	15.3	1.6	1.6
	percent	.022	.019	.051	.016	.016
IV	seconds	0.3	0.0	0.0	0.0	1.1
	percent	.001	.000	.000	.000	.002
V	seconds	0.7	2.4	0.6	2.4	1.7
	percent	.002	.008	.002	.005	.004

Inspection of the graph for Child I indicates that during Phase I no consistent or significant effects results from presentation of the independent variable. However during Phase II, five times as much off-task behavior was associated with the more difficult of the two levels.

Child II displayed the greatest amount of off-task behavior overall. On most experimental days, presentation of the higher difficulty level tasks resulted in proportionately higher amounts of off-task behavior.

Data for Child III during Phase I indicate that 2 1/2 times as much off-task behavior were associated with level 3 as compared to level 1. However, when this subject was presented two tasks, one easily within her ability repertoire and the other exceptionally difficult, she exhibited approximately equal amounts of off-task behavior on both.

Table 2 presents ratios that permit direct comparison of the size of the experimental effect. Comparisons are made between various levels within phases. Critical ratios are the 3:1 and 4:1 comparisons in which the most difficult levels are compared with the easiest level. Off-task behavior for Child II doubled in the 3:1 comparison. This change represents an increase of 20 percent over baseline level (Table 1). While some data for Subjects I and III indicate higher multiples of change, the greatest amount of practical time change was shown by Child II.

Presentation Order

To determine if order of task presentation affected the dependent variable, percentages of time off-task per randomized presentation were computed. These percentages are offered in Table 3. No order by level of difficulty interactions are apparent.

Table 2
Ratio Comparison of Mean Seconds
Off-Task Between Levels

Subject	Level			
	2:1	Phase I 3:1	3:2	Phase II 4:1
I	0.89	1.15	1.29	5.30
II	1.61	2.00	1.24	--
III	.87	2.28	2.64	1.00
IV	.30	.00	.00	1.10
V	—*	.86	.25	.71

*Division for this data point yields a large ratio of 3.43. This ratio is misleading since the actual values involved (2.4 secs. and .7 secs.) are too small to interpret meaningfully.

Table 3
Mean Percentage of Time Off-Task
by Order of Presentation

Subject	Order of Presentation				
	1st	Phase I 2nd	3rd	Phase II 1st	2nd
I	.081	.077	.082	.086	.049
II	.328	.305	.359	--	--
III	.036	.038	.022	.026	.006
IV	.001	.000	.000	.002	.000
V	.006	.004	.002	.008	.001

Interest/Satisfaction Rating

Scaled responses to the question, "How did you like the story you just read?" were recorded after individual reading sessions. These results, expressed as percentages, are presented in Table 4.

In most instances the subjects indicated that they did not like the more difficult reading material but enjoyed the easy material. However, there were two subjects (Children I and V) who rated all the levels as generally favorable.

Correlates of Off-Task Behavior

Appendix B provides a description of subjects participating in this study. A visual inspection of background data, including age, sex, ethnic origin, reading group, and IQ score, shows that the subjects do not share specific characteristics linking them to greater or lesser durations of off-task behavior. The child exhibiting the greatest amount of off-task behavior in situations of mismatch (II) was of the same race and reading group and very close to the same age and IQ as another subject (V) showing practically no off-task behavior whatever.

Similar observations are made concerning on-task durations. For the two subjects with minimal off-task behavior (IV and V), IQ scores were 98 and 77, one was male and the other female, one was white and the other black, and 1 year 7 months separated them in age. Sufficient discrepancies exist, therefore, so that no conclusions can be drawn as to whether the above mentioned demographic,

Table 4
Mean Percentages for Satisfaction Ratings
by Difficulty Level

Subject	Level		Ratings*			
	Phase I	Phase II	1	2	3	4
I	1		.27	.64	.00	.09
	2		.55	.45	.00	.00
	3		.36	.64	.00	.00
		1	.00	1.00	.00	.00
		4	.00	.78	.22	.00
II	1		.00	.67	.33	.00
	2		.00	.61	.22	.17
	3		.06	.39	.33	.22
III	1		.87	.13	.00	.00
	2		.67	.00	.20	.13
	3		.60	.07	.26	.07
		1	1.00	.00	.00	.00
		4	.25	.00	.25	.50
IV	1		1.00	.00	.00	.00
	2		.12	.25	.00	.63
	3		.37	.00	.00	.63
		1	1.00	.00	.00	.00
		4	.29	.00	.00	.71
V	1		1.00	.00	.00	.00
	2		1.00	.00	.00	.00
	3		1.00	.00	.00	.00
		1	.86	.14	.00	.00
		4	.29	.57	.00	.14

*1 - like a whole lot

2 - liked a little

3 - did not like very much

4 - did not like at all

intellectual, and academic factors are critical variables affecting off-task behavior in situations of mismatch.

Summary

In summary, the results show that off-task behavior in three of five experimental subjects was related to the difficulty level of academic tasks presented them. For each case in which an experimental effect was noted, off-task behavior increased as difficulty level increased. For the remaining two subjects, practically no experimental effects whatsoever were measured. Possible explanations for this finding will be offered in the next chapter.

With regard to order of task presentation, no evidence was found to indicate that any particular place in the order or sequence of presentation was associated with more or less off-task behavior. An additional finding was that a majority of the subjects expressed greater dissatisfaction with the more difficult task material. While this was the general situation, exceptions were also noted and some children expressed the same basic sentiments for all levels. Finally, among such known subject variables as age, sex, ethnic origin, IQ, and reading group, no relationships between them and off-task behavior were observed.

CHAPTER V

DISCUSSION

This study sought to provide empirical data which would deal directly with a popularly-held educational assumption, namely that students suffer adversely when using educational materials ill-suited to their ability (Durrell, 1940; Harris & Sipay, 1975). In situations of mismatch, children are frequently described as becoming "confused," "anxious," "frustrated," and "bored" (Durrell, 1940; May, 1973). Instead of attempting to define and measure intangible and speculative concepts such as "frustration" or "anxiety," this study used direct measurement techniques to record clearly observable behaviors. These behaviors, which may or may not be correlates of such constructs as frustration and anxiety are themselves the subject of study.

This chapter will begin with a discussion of the findings. Following this will be a discussion of problems and limitations encountered in this study. Finally, conclusions, implications, and suggestions for future research will be addressed.

Findings

The stated purpose of this study was to investigate the relationship between varying levels of academic task difficulty and

off-task classroom behaviors. In order to meet this objective the study had to show that data for a single subject in an intensive experimental design provided a clear and systematic experimental effect. The design used allows conclusive statements of findings when these findings are consistent and pronounced within a single subject. Additional replications of the experiment with other subjects may be sought to increase the soundness and generalizability of these findings. However, each experiment retains its individual integrity. Two separate experiments, each involving a single subject, provide clear indication that task difficulty does indeed effect overt classroom behavior in a systematic way. The effect is also present in a third subject but the overall durations involved are quite small.

A consistent trend in the data was found for Child II. In general, presentation of progressively more difficult reading tasks resulted in progressive increases in duration of off-task behavior. On the average, Child II spent over 40 percent of level 3 task sessions exhibiting off-task behaviors. This figure is more than double the time off-task for level 1 sessions. At one point, off-task behavior during level 3 sessions rose to 75 percent of the five minute reading period.

Observable behavioral changes quickly began to occur during experimental sessions. As material difficulty increased, Child II increased or initiated such behaviors as talking to his neighbors, glancing at the observer and about the room, and throwing objects such as erasers or paper. He would also hold his book up before his face, peer over the edge, and look about.

Child I provides additional evidence of an experimental effect. Although this subject persevered on task throughout Phase I, he was virtually driven off-task by the introduction of an exceedingly difficult reading task in Phase II. As a result, five times as much off-task behavior is associated with the more difficult reading level during Phase II.

Data for Child I, while corroborating the effect for Child II, also provide an interesting contrast in the ability of children to tolerate difficult tasks. While Child II was easily influenced by progressively more difficult tasks, Child I was more resistant. This subject, who was drawn from a higher reading group than Child II, seemed to meet the challenge of the more difficult Phase I tasks. Better readers, such as Child I, having a longer history of success on similar academic tasks may persevere longer, albeit with increased effort, in an attempt to overcome difficulties.

Results of the experiment with Child II strongly suggest that systematic changes in academic task difficulty can result in systematic changes in off-task behavior for an elementary school child. This finding coincides with conclusions drawn by Jorgenson (1977, 1978) and Jorgenson, Klein, and Kumar (1977) that as reading material becomes more difficult for a student, the student's behavior (as judged by the teacher) becomes less appropriate for the classroom.

While this study was successful in finding evidence that a mismatch of task difficulty and ability results in an increase of inappropriate classroom behaviors, there are also indications that under certain conditions task difficulty has little or no effect on

the off-task behavior of third grade children. This latter finding is based on data generated by Subjects IV and V and to a lesser extent by Subject III.

These three children maintained almost total attention to task despite the extreme difficulty of certain reading passages. There were, however, certain exceptional behaviors noted during experimental conditions with these subjects which deserve mention. These behaviors, occurring during presentation of the more difficult levels, included long periods of eye rubbing, quick glances about the room (with durations of less than two seconds), rapid page turning, yawning, grimacing, sighing, and fidgeting. Since these subjects manifested these additional behavioral signs while maintaining eye contact with the written page, no off-task behavior was recorded.

An additional behavior occurring at the more difficult levels was balking at the prospect of further reading. On those occasions when a subject balked, gentle encouragement was successful in persuading him or her to continue. This latter occurrence is indicative of the ease with which these subjects were influenced by social reinforcers.

The notion that social reinforcers may have served to maintain or increase attention-to-task behavior deserves additional comment. Several of the children, namely III, IV, and V, were especially agreeable and cooperative subjects and seemed eager to please the experimenter. Child II and to a lesser extent Child I were more reticent. Gewirtz and Baer (1958) found that adult verbal approval had definite reinforcing value for children and that the effectiveness

of adult approval as a reinforcer increased after periods in which social reinforcement was not available. In this study, adult observers were required to present each of three reading tasks to the subjects, then leave them and begin observations. While conversation was held to a minimum, the observer did have to obtain a satisfaction rating and give instructions on where to begin the next reader. For some subjects, especially those with a strong desire to cooperate (or need for approval) these encounters with an adult may have reinforced on-task behavior. The conditions under which subjects in this study received social reinforcement in other settings and the amount which they received is unknown. However, it is Gerwitz and Baer's (1958) contention that the amount of reinforcing power gained by social reinforcers is directly related to the degree to which an individual child seeks such approval in other settings.

There are other factors besides social reinforcement which may have been working to sustain on-task behavior. The desire to do well and learning to follow directions given by adults can certainly be largely responsible. Certain other factors are more subtle. For example, during the course of this study it was noticed that after completion of the experimental sessions for the day, subjects III, IV, and V would at times talk and joke with their neighbors or the observer and walk about the room, all of which were behaviors inappropriate in the classroom at the time. Those children may have felt considerable desire and/or pressure to conform for the sake of a project which they felt an important part of and became more regular in their behavior only when the pressure

to comply was off. Thus, the "guinea pig effect" (Isaac & Michael, 1971) in which the acute awareness a subject feels for an experimental situation and the desire to perform in a way that is believed satisfactory for the experiment, may have been a factor serving to suppress more normally occurring off-task behavior.

It may also be more than coincidence that Subjects III, IV, and V came from the same classroom while I and II came from another. Although only based on casual observations it appeared that one teacher was more authoritative in her interactions with the children. This teacher was quite intolerant of even low noise levels in her classroom and frequently threatened harsh punishment for infractions of the rules. In contrast, the other teacher was soft-spoken, permissive, and rarely made overt threats to the children. Subjects I and II were drawn from this latter teacher's classroom, while III, IV, and V came from the former. It has been demonstrated with both animals and humans that aversive control generalizes across behaviors, settings, and through time (Azrin & Holz, 1966). Whether such variables as teacher's style and personality are significant factors contributing to off-task behavior in situations of mismatch is a subject for future study.

Problems and Limitations of the Study

One of the foremost limitations to a study of children in a natural classroom setting is the reactive influence of the study's procedural methodology. According to Isaac and Michael (1971) "a measurement is reactive whenever the subject is directly involved in a study and he is reacting to the measurement process itself."

In this study, reactive measurement may have assumed two forms. One is the "guinea pig effect" mentioned earlier. With this form the subject feels he must do his best for the experiment and therefore becomes unusually cooperative. Thus the act of participating in the process itself changes the subject in ways which in turn affect the research results.

The second form of reactive measures relevant to this study is that of role selection. In this case, an individual chosen as a subject in the study "assumes a role of" the kind of person "he thinks he should be" (Isaac & Michael, 1971). To the degree that the subject acts in a way that is different from the way that he or she would ordinarily act, the research results are affected. Participation in a "special project" and the inevitable realization that one is being specifically observed seems to have had marked effects on certain subjects in this study. However, circumventing such a set of circumstances is usually very difficult. The child's permission to participate must be obtained as must his cooperation during the task sessions and unless one-way observation windows are available, the child soon realizes he is being watched.

Two problems were found to exist when using a school adopted basal reader series. It was found that children who were fluent readers performed almost equally as well on all three upper levels of the Ginn Series. In other words, for these children a functional difference in difficulty level did not seem to exist at the higher levels. In order to provide another level of difficulty with sufficient functional differentiation, it was necessary to select a text that was not a part of the series.

The second problem was that readers in the Ginn Series were found to lack uniform difficulty throughout. While three 100 word reading samples were selected from the beginning, middle, and end of each reader and performance on these then averaged, fluctuations in difficulty level still remained within a particular reader from session to session. While a problem in one sense, it is an asset in another since this experimental situation more closely approximates real-life classroom situations.

Conclusions and Implications

It has been the thesis of this study that academic task difficulty is a powerful factor influencing the everyday classroom behavior of pupils. A major conclusion drawn from the evidence is that a direct relationship between academic task difficulty and off-task behaviors does exist for some subjects. In general, as a reading task becomes increasingly difficult, some children become increasingly non-attentive and (to a certain extent) disruptive within the classroom. This conclusion supports a long standing common sense assumption concerning the effects of task difficulty on behavior and corroborates the findings of Jorgenson (1977, 1978) and Jorgenson, Klein, and Kumar (1977).

While the evidence is clear and convincing that task difficulty does affect classroom behavior in an observable way, this study also presents evidence that some children are not so affected. Therefore, the theoretical assumption that "a learning task which is too easy for children will be boring, and one that is too hard will cause them to withdraw" (May, 1973) was not wholly supported. A

differential reaction to task difficulty within an experimental setting appears to be the case, and experimenter effects may play a significant role.

Conclusions drawn from this study afford a number of practical implications for those involved in research or applied practice. First, the results provide a base for considering real environmental rather than hypothetical or organic factors as significant contributors to learning disabilities and emotional disturbance in school children. In fact, many children labeled specific learning disabled (SLD) or emotionally handicapped (EH) would perhaps not qualify for these special education programs if careful analysis was made of the relationship between a student's performance level and the tasks required of him. Smith and Neisworth (1975) caution against labeling because we are all "gifted, average, or slow according to the task at hand." Thus, attentional deficits, learning problems, day-dreaming, and disruptive behavior are linked to processes operating within the classroom. The problem is now approached as an interaction of child and learning environment.

Second, the evidence indicates to classroom teachers that a first concern should be the careful assessment of performance and placement in work assignments commensurate with that performance. Student performance may in large measure be due to the procedures used to establish that performance behavior. Frequent functional assessment helps to establish and maintain an accurate match of performance and task difficulty. Reliance on direct measurement procedures involving rate and comprehension percentage circumvents the problems of inaccuracies in publisher designated grade levels

(Britton & Lumpkin, 1977; Pruitt, 1977), unreliable reliability formulae (Bormuth, 1966), and the once-a-year standardized test ritual. Materials actually available for use with the child can be used in a curriculum-based assessment and then matched to the child's current performance level. In this way the teacher is able to avoid a significant source of error in measurement--the inferential transition between standardized test score and actual task material.

Commensurate with the need for frequent direct assessment is acceptance on the part of the teacher of a pupil's current level of performance. Well-meaning teachers seeking to provide a challenge for students may not be achieving this goal. According to Zintz (1972), "too many times the child is 'pushed' to work at his frustration level rather than his instructional level." Thus, a mismatch may foster negative attitudes toward the subject, the teacher, and even the school itself.

A final implication of this study concerns the utility of considering many behavior problems as classroom based. Prospects for remediation are greatly increased when experimental effects are established to be the result of factors within the learning environment. Usually these factors can be effectively influenced by the classroom teacher. Certainly performance level/difficulty level matching can be. However, when classroom problems are initially believed to be organically derived, a teacher's chances for effecting behavioral change are considerably reduced. Dealing with organic problems, to include varying psychiatric syndromes of emotional disturbance, usually requires professionals who are not school based.

Thus, the problem has left the teacher's sphere of control. In order to have solutions to off-task behavior problems reenter the domain of educators, practices such as performance/task matching must be examined as a routine first order of business. The notion of organicity should be put off until after these careful assessments are made. Great savings of time, expense, and wear and tear on children would result from seeking the simple solution first.

Suggestions for Further Research

There are an unlimited number of questions generated by a research endeavor, especially one having had a paucity of research attention. Replicating a study provides a reliable way of confirming or disclaiming previous findings (Sidman, 1960). Sidman (1960) has discussed two types of replication. Direct replication, in which all major experimental variables remain the same, is unrealistic to expect in natural settings. Systematic replication, in which experimental variables are systematically changed, does provide promise for applied research. In replicating this study, several suggestions may help to generate clear and convincing findings. These include the following:

1. Utilize a laboratory school or other school setting with special observation facilities. The use of a one-way observation mirror would have provided a more unobtrusive measure and videotaping equipment permits a careful review of recorded behaviors.
2. Systematically replicate the experiment with children identified as discipline problems in the regular classroom or with

children labeled EH or SLD. These children may show very dramatic changes in off-task behavior when subjected to difficult tasks.

3. Select subjects of different age levels to determine if age is a factor in the treatment effect.

4. Obtain baseline measures of off-task behavior before any contact with the child is made. Children could be observed during regular class academic assignments as long as a certain period of the class day stipulates silent reading or quiet seatwork. Such a baseline would more accurately reflect a child's normal level of off-task behavior and allow more accurate conclusions to be drawn about a treatment effect. Reactive effects would be more likely if children showed the same or a lower level of off-task behavior during treatment.

5. Define, observe, and record off-task behavior after experimental sessions are completed. Compare these data with the baseline to determine any "after effects" of treatment.

6. In order to be able to generalize across tasks, arithmetic tasks could be substituted for reading tasks. A functional method of assessing difficulty level of arithmetic problems could be obtained or devised.

7. Investigate the variable of authoritative versus permissive teaching style as a factor affecting off-task behavior. The number of reprimands or threats of punishment used by different teachers participating in the study can be counted. Likewise a count can be made of the number of instances teachers use certain positive reinforcers. Inhibition or expression of off-task behavior could then be compared to the frequency with which teachers use positive or negative attention.

REFERENCES

- Arnold, R. D., & Sherry, N. A comparison of the reading levels of disabled readers with assigned textbooks. Reading Improvement, 1975, 12, 207-211.
- Auvenshine, A. L. B. A study of the readability of junior and community college textbooks used in the academic areas and the reading abilities of students using the textbooks (Doctoral dissertation, Baylor University, 1978). Dissertation Abstracts International, 1978, 39, 2710A.
- Ayllon, T., & Roberts, M. D. Eliminating discipline problems by strengthening academic performance. Journal of Applied Behavior Analysis, 1974, 7, 71-76.
- Azrin, N. H., & Holz, W. C. Punishment. In W. K. Honig (Ed.), Operant behavior: Areas of research and application. Englewood Cliffs, N.J.: Prentice-Hall, 1966, 380-447.
- Bacon, H. K., Child, I. L., & Barry, H. A. A cross-cultural study of correlates of crime. Journal of Abnormal and Social Psychology, 1963, 66, 291-300.
- Bailey, J. S. A handbook of research methods in applied behavior analysis. Copyright 1977, in press.
- Baldwin, T. L., Johnson, T. J., & Wiley, D. E. The teacher's perception and attribution of causation. Paper presented at the meeting of the American Educational Research Association, Minneapolis, April, 1970.
- Bankston, L. V. C. The effects of the readability of mathematical materials on achievement in remedial mathematics in a selected community college (Doctoral dissertation, University of Texas at Austin, 1975). Dissertation Abstracts International, 1975, 36, 815A.
- Becker, W. C. Consequences of different kinds of parental discipline. In M. L. Hoffman & L. W. Hoffman (Eds.), Review of child development research (Vol. 1). New York: Russel Sage, 1964, 169-208.

- Becker, W. C., Madsen, C. H., Arnold, C. R., & Thomas, D. R. The contingent use of teacher attention and praise in reducing classroom behavior problems. The Journal of Special Education, 1967, 1, 287-307.
- Bertalan, J. J. A comparison of the readability of community college social science textbooks with student reading levels and the effect of academic achievement (Doctoral dissertation, Florida Atlantic University, 1977). Dissertation Abstracts International, 1977, 38, 4596A.
- Betts, E. A. Foundations of reading instruction. New York: American Book Company, 1946.
- Bond, G. L., & Dykstra, R. The cooperative research program in first grade reading instruction. Reading Research Quarterly, 1967, 2, 5-142.
- Bond, G., & Tinker, M. Reading difficulties: Their diagnosis and correction. New York: Appleton-Century-Crofts, 1973.
- Bormuth, J. R. Readability: A new approach. Reading Research Quarterly, 1966, 1, 79-132.
- Bradley, J. M., & Ames, W. S. Readability parameters of basal readers. Journal of Reading Behavior, 1977, 9, 175-183.
- Britton, G. E., & Lumpkin, M. C. Computerized readability verification of textbook reading levels. Reading Improvement, 1977, 14, 193-199.
- Broden, M., Bruce, C., Mitchell, M., Carter, V., & Hall, R. V. Effects of teacher attention on attending behavior of two boys at adjacent desks. Journal of Applied Behavior Analysis, 1970, 3, 199-203.
- Bucher, B., & Okovita, H. W. Effects of differential reinforcement and task difficulty on preschoolers compliance with teacher instructions. Journal of Experimental Child Psychology, 1977, 23, 226-236.
- Camp, B., & Zimet, S. Classroom behavior during reading instruction. Exceptional Children, 1975, 42, 109-110.
- Campbell, B. R. The relationship of reading ability and comprehension of social studies and science textbook selection (Doctoral dissertation, Florida State University, 1974). Dissertation Abstracts International, 1973, 33, 3146A.

- Carkhuff, R. R., & Berenson, B. G. Beyond counseling and therapy (2nd ed.). New York: Holt, Rinehart and Winston, 1977.
- Chall, J. S. Learning to read: The great debate. New York: McGraw-Hill, 1967.
- Chall, J. S., & Feldman, S. First grade reading: An analysis of the interactions of professed methods, teacher implementation, and child background. Reading Teacher, 1966, 19, 569-575.
- Clymer, T., and others. Reading 720. Lexington, Mass.: Ginn, 1976.
- Cobb, J. A. Relationship of discrete classroom behaviors to fourth grade academic achievement. Journal of Educational Psychology, 1972, 63, 74-80.
- Drillien, C. The incidence of mental and physical handicaps in school age children of very low birth weight. Pediatrics, 1961, 27, 452.
- Duke, D. L. Adults can be discipline problems too! Psychology in the Schools, 1978, 15, 522-28.
- Durrell, D. Improvement of basic reading abilities. Yonkers, N.Y.: World Book, 1940.
- Eberwein, L. D. The variability of readability of basal reader textbooks and how much teacher know about it. Reading World, 1979, 18, 259-272.
- Ewing, M. J. A comparison of the effects of readability and time on learning the content of a state driver's handbook (Doctoral dissertation, Florida State University, 1976). Dissertation Abstracts International, 1976, 37, 6154A.
- Feldhusen, J. F., Thurston, J. R., & Benning, J. J. A longitudinal study of delinquency and other aspects of children's behavior. International Journal of Criminology and Penology, 1973, 1, 341-351.
- Fendrick, P. A., & Bond, G. L. Delinquency and reading. Pedagogical Seminary and Journal of Genetic Psychology, 1936, 48, 236-243.
- Fox, D. L. The reading achievement levels of college freshmen enrolled in selected first-semester college courses compared with the readability levels of textbooks assigned in those courses (Doctoral dissertation, East Texas State University, 1978). Dissertation Abstracts International, 1978, 39, 2721A.

- Fry, E. The readability principle. Language Arts, 1975, 52, 847-851.
- Gates, A. I. Maladjustment due to failure in reading. School Executive, 1933, 55, 379-380.
- Gates, A. I. The improvement of reading (3rd ed.). New York: Macmillan, 1947.
- Gerwurtz, J., & Baer, D. The effect of brief social deprivation on behaviors for a social reinforcer. Journal of Abnormal Social Psychology, 1958, 56, 49-56.
- Gilmore, G. E., & Chandy, S. M. Educators describe the school psychologist. Psychology in the Schools, 1973, 10, 397-403.
- Glavin, J. P., & Annesley, F. R. Reading and arithmetic correlates of conduct/problem and withdrawn children. Journal of Special Education, 1971, 5, 213-219.
- Goodlad, J. I. School, curriculum and the individual. New York: John Wiley & Sons, 1966.
- Granzow, K. R. A comparative study of underachiever, normal achiever, and overachiever in reading (Doctoral dissertation, State University of Iowa, 1954). Dissertation Abstracts, 1954, 14, 631-632.
- Gronlund, N. E. Measurement and evaluating in teaching (3rd ed.). New York: Macmillan, 1972.
- Haring, N. G., Lovitt, T. C., Eaton, M. D., & Hansen, C. L. The fourth R: Research in the classroom. Columbus, Ohio: Charles E. Merrill, 1978.
- Harris, A. J., & Sipay, E. R. How to increase reading ability. New York: David McKay, 1975.
- Isaac, S., & Michael, W. B. Handbook in research and evaluation. San Diego, Calif.: Edits, 1971.
- James, W. Principles of psychology. New York: Holt, 1890.
- Jessor, J., & Jessor, S. L. Problem behavior and psychosocial development. New York: Academic Press, 1977.
- Jones, D. M. An experiment in adaptation to individual differences. The Journal of Educational Psychology, 1948, 39, 257-272.
- Jorgenson, G. W. An analysis of teacher judgments of reading level. American Educational Research Journal, 1975, 12, 67-75.

- Jorgenson, G. W. Relationship of classroom behavior to the accuracy of the match between material difficulty and student ability. Journal of Educational Psychology, 1977, 69, 24-32.
- Jorgenson, G. W. Student ability--material difficulty matching: Relationship to classroom behavior. Paper presented at the American Educational Research Association Conference, Toronto, Canada, 1978.
- Jorgenson, G. W., Klein, N., & Kumar, V. K. Achievement and behavior correlates of matched levels of student ability and material difficulty. Journal of Educational Research, 1977, 71, 100-103.
- Karlsen, B. Comparison of some educational and psychological characteristics of successful and unsuccessful readers at the elementary school level (Doctoral dissertation, University of Minnesota, 1954). Dissertation Abstracts, 1955, 15, 456A.
- Keogh, B. K., Welles, M. F., & Weiss, A. L. Field independence-dependence, reflection-impulsivity, and problem solving styles of preschool children. Technical Report SERPA-1. Los Angeles: University of California, 1972.
- Klare, G. R. The measurement of readability. Iowa State University, 1963.
- Lahaderne, H. M. Attitudinal and intellectual correlates of attention: A study of four sixth-grade classrooms. Journal of Educational Psychology, 1968, 59, 320-324.
- Lovitt, T. C., & Hansen, C. L. Round one--placing the child in the right reader. Journal of Learning Disabilities, 1976, 9, 347-359.
- May, F. To help children read. Columbus, Ohio: Charles E. Merrill, 1973.
- Miller, J. W., & Marshall, F. W. Teacher's abilities to judge the difficulty of reading materials. Paper presented at the Annual Meeting of the IRA. Houston, Texas, May 1-5, 1978.
- Mills, R. E., & Richardson, J. R. What do publishers mean by "grade level?" The Reading Teacher, 1963, 16, 359-362.
- Moyer, K., & Gilmer, B. E. The concept of attention spans in children. Elementary School Journal, 1954, 54, 464-466.
- National Education Association. Teacher opinion poll: Discipline. Today's Education, 1976, 65, 20.

- Nottelman, E. D., & Hill, K. T. Test anxiety and off-task behavior in evaluative situations. Child Development, 1977, 48, 225-231.
- Paige, W. D. The effects of lowered readability and the use of a glossary upon reading comprehension of selected industrial arts electronics text materials (Doctoral dissertation, Ohio State University, 1978). Dissertation Abstracts International, 1978, 39, 2792A.
- Pruitt, J. L. The readability of selected fourth, fifth, and sixth grade social studies textbooks as determined by the Fry readability graph and group informal reading inventories (Doctoral dissertation, Oregon State University, 1977). Dissertation Abstracts International, 1977, 38, 2529A.
- Robeck, M., & Wilson, J. Psychology of reading. New York: John Wiley and Sons, 1974.
- Ruble, D. M., & Nakamura, C. Y. Task orientation versus social orientation in young children and their attention to relevant social cues. Child Development, 1972, 43, 471-480.
- Samuels, S. J., & Turnure, J. E. Attention and reading achievement in first grade boys and girls. Journal of Educational Psychology, 1971, 66, 29-32.
- Sidman, M. Tactics of scientific research: Evaluating experimental data in psychology. New York: Basic Books, 1960.
- Skinner, B. F. Technology of teaching. New York: Appleton-Century-Crofts, 1968.
- Smith, R. M., & Neisworth, J. T. The exceptional child. New York: McGraw-Hill, 1975.
- Spache, G. D. Personality patterns of retarded readers. Journal of Educational Research, 1957, 40, 461-469.
- Staiger, R. C. Basal reading programs: How do they stand today? In N. B. Smith (Ed.), Current issues in reading. Newark, Del.: International Reading Association, 1969.
- Stradley, W. E., & Aspinall, R. D. Discipline in the junior high/middle school. New York: The Center for Applied Research in Education, 1975.
- Swift, M., & Spivack, G. Alternative teaching strategies: Helping behaviorally troubled children achieve. Champaign, Ill.: Research Press, 1975.

- Symonds, P. M. Diagnosing personality and conduct. New York: Century, 1931.
- Thomas, D. R., Becker, W. C., & Armstrong, M. Production and elimination of disruptive classroom behavior by systematically varying teacher's behavior. Journal of Applied Behavior Analysis, 1968, 1, 35-45.
- Thorndike, E. L. Fundamentals of learning. New York: Teachers' College, Columbia University, 1932.
- Thorndike, R. L., & Hagen, E. Measurement and evaluation in psychology and education (3rd ed.). New York: Wiley, 1969.
- Tinker, M., & McCullough, C. Teaching elementary reading. Englewood Cliffs, N.J.: Prentice-Hall, 1975.
- Tukey, J. W. Exploratory data analysis. Reading, Mass: Addison-Wesley, 1977.
- Zintz, M. V. Corrective reading (2nd ed.). Dubuque, Iowa: W. C. Brown, 1972.

APPENDIX A
DEFINITION OF TERMS

Difficulty Level: one of three levels corresponding to criteria for correct words per minute, incorrect words per minute, and comprehension percentage.

Off-Task Behavior: (1) looking away from assigned printed material for two seconds or more; (2) closing eyes for two seconds or more.

On-Task Behavior: in seat with eyes focused on work for at least two seconds.

Readability: the relative difficulty of a piece of printed material.

APPENDIX B
DESCRIPTION OF SUBJECTS

Subject	Sex		Ethnic Origin	Olat Score*	Reading Group
I	M	8 yrs. 0 mos.	W	--	Average
II	M	9 yrs. 10 mos.	B	87	Below Average
III	F	7 yrs. 11 mos.	H ⁺	128	Average
IV	M	8 yrs. 1 mo.	W	98	Below Average
V	F	9 yrs. 8 mos.	B	77	Below Average

*Otis-Lennon Mental Abilities Test administered 4/79.

⁺Hispanic

APPENDIX C
RAW AND SMOOTHED DATA

Child I

Session #	1		2		3		4	
	Raw	Smoothed	Raw	Smoothed	Raw	Smoothed	Raw	Smoothed
1	3	6	18	12	5	3		
2	9	9	5	5	0	5		
3	36	9	0	5	34	34		
4	4	36	31	31	37	34		
5	60	60	43	31	25	37		
6	85	60	20	20	58	25		
7	0	0	0	20	0	18		
8	0	0	24	24	18	18		
9	6	4	29	25	53	53		
10	4	6	25	29	62	53		
11	54	4	37	31	8	35		
12	0	0	-	-	-	-	23	13
13	0	0	-	-	-	-	2	20
14	5	5	-	-	-	-	20	20
15	26	23	-	-	-	-	99	99
16	23	23	-	-	-	-	105	99
17	25	23	-	-	-	-	28	97
18	0	8	-	-	-	-	97	28
19	8	0	-	-	-	-	21	68
20	0	4	-	-	-	-	68	45
Totals	(348)	(280)	(232)	(253)	(300)	(315)	(463)	(489)

Child II

Session #	Seconds Off-Task By Level					
	Raw	1 Smoothed	Raw	2 Smoothed	Raw	3 Smoothed
1	1	4	26	13	132	85
2	6	1	0	26	37	69
3	0	6	145	94	69	69
4	75	13	94	94	206	69
5	13	17	86	86	61	74
6	17	17	20	86	74	74
7	91	91	155	149	146	89
8	212	91	149	149	89	89
9	55	72	40	50	72	89
10	72	72	50	50	263	147
11	110	72	115	103	147	208
12	33	102	103	103	208	147
13	102	52	93	103	125	154
14	52	102	207	207	154	154
15	134	52	210	207	181	154
16	31	31	79	128	104	104
17	12	31	128	79	79	104
18	80	46	70	94	141	110
Totals	(1096)	(872)	(1770)	(1821)	(2288)	(1919)

Child III

Session #	Seconds Off-Task by Level							
	1		2		3		4	
	Raw	Smoothed	Raw	Smoothed	Raw	Smoothed	Raw	Smoothed
1	0	0	11	6	14	15		
2	0	0	0	0	15	15		
3	3	0	0	0	21	15		
4	0	3	15	1	0	21		
5	15	15	1	15	90	19		
6	21	15	30	1	19	19		
7	0	0	0	0	10	10		
8	0	0	0	0	0	10		
9	0	0	0	0	29	0		
10	0	0	20	10	0	24		
11	6	0	10	10	24	0		
12	0	6	0	0	0	8		
13	55	0	0	0	8	0		
14	0	0	0	0	0	0		
15	0	0	1	1	0	0		
16	0	0	-	-	-	-	5	5
17	0	0	-	-	-	-	4	4
18	15	0	-	-	-	-	0	4
19	0	0	-	-	-	-	33	0
20	0	0	-	-	-	-	0	0
21	7	0	-	-	-	-	0	0
22	0	7	-	-	-	-	0	0
23	35	18	-	-	-	-	17	9
Totals	(154)	(64)	(88)	(44)	(220)	(156)	(59)	(22)

Child IV

Session #	Seconds Off-Task By Level							
	1		2		3		4	
	Raw	Smoothed	Raw	Smoothed	Raw	Smoothed	Raw	Smoothed
1	0		0		0			
2	0		0		0			
3	0		0		0			
4	0		0		0			
5	0		0		0			
6	2		0		0			
7	0		0		0			
8	0		0		0			
9	0						5	
10	0						3	
11	0						0	
12	0						0	
13	0						0	
14	0						0	
15	0						0	
Totals	(2)		(0)		(0)		(8)	

Child V

Session #	Seconds Off-Task By Level							
	1		2		3		4	
	Raw	Smoothed	Raw	Smoothed	Raw	Smoothed	Raw	Smoothed
1	0		6		0			
2	0		0		0			
3	0		3		0			
4	0		2		0			
5	0		0		6			
6	5		0		0			
7	0		0		0			
8	0		0		0			
9	2		13		0			
10	0		0		0			
11	0						0	
12	0						9	
13	4						0	
14	0						3	
15	0						0	
16	0						0	
17	13						0	
Totals	(24)		(21)		(6)		(12)	


BIOGRAPHICAL SKETCH

Paul J. Siegel received his B.A. in psychology from the University of Florida in March of 1969. Travel in Europe and South America followed. A three year military career then ensued in which Mr. Siegel attended Officer Candidate School and served in Viet Nam. Discharged as a 1st Lieutenant, he returned to the University of Florida to pursue a career in school psychology.

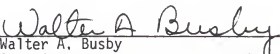
In 1978, Mr. Siegel received his M.Ed. and Ed.S. degrees in school psychology and began pursuit of the Ph.D. degree. He served as graduate assistant for the school psychology program from 1977-1979.

Mr. Siegel is a member of the Florida Association of School Psychologists and National Association of School Psychologists and is currently employed as a school psychologist in Clay County, Florida.

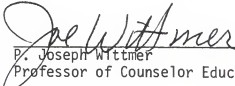
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


William D. Wolking, Chairman
Professor of Counselor Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Walter A. Busby
Associate Professor of Foundations
of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


P. Joseph Wittmer
Professor of Counselor Education

This dissertation was submitted to the Graduate Faculty of the Department of Counselor Education in the College of Education and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1980

Dean, Graduate School